Complex Falk-Langemeyer Method

Vjeran Hari[1,*]

[1] Department of Mathematics, University of Zagreb, Bijenička cesta 30, 10000 Zagreb, Croatia

The known Falk-Langemeyer method for the simultaneous diagonalization of two positive definite symmetric matrices is generalized to work with complex matrices. It is shown that the derived method is well defined for the Hermitian matrices which make a definite pair. Special attention is paid to the stability of the formulas for the computational parameters when the pivot submatrices are close to being proportional. Numerical tests indicate the high relative accuracy of the method provided that both matrices are definite and well-behaved, i.e. if they can be well-scaled symmetrically.

Copyright line will be provided by the publisher

1 Introduction

In 1960 Falk and Langemeyer [1] proposed a Jacobi-type method for solving the generalized eigenvalue problem (GEP) $Ax = \lambda Bx$, $x \neq 0$ with symmetric positive definite matrices A, B. Later Slapničar and Hari [8] proved the quadratic convergence of the method in the case of simple eigenvalues. They also proved that the method was well defined for a more general, definite pair of symmetric matrices. In 2015 Matejaš [6] proved sharp error estimates for the method. Our numerical tests indicate the high relative accuracy of the method in the case of positive definite matrices that can be well-scaled symmetrically, i.e. when the condition numbers of $D_A A D_A$ and $D_B B D_B$ are small for some diagonal matrices D_A and D_B. Since it is a Jacobi method for the GEP it is very efficient and highly accurate when both matrices are almost diagonal (cf. [5]). This typically happens in the course of modelling the parameters of a system. The described properties make the method an excellent choice for the kernel algorithm of the block-Jacobi methods which are nowadays the prime choice of the methods for solving the definite GEP on contemporary parallel CPU and GPU computing machines [7].

In this short communication, we present the main formulas of the complex Falk-Langemeyer (CFL) algorithm that are derived in [3]. Although they are the proper generalization of those in the real method, their derivation is not trivial. The formulas for the transformation parameters become useless when the pivot submatrices are proportional or very close to being proportional. In that case, several stable formulas are proposed in [3]. It is proved that the new complex algorithm is well defined when some real linear combination of the complex Hermitian matrices A and B is positive definite. Numerical tests strongly indicate the high relative accuracy of the method when both matrices are positive definite and can be well-scaled symmetrically. The quadratic convergence of the method can be proved following the original proof from [8]. The proof of the global convergence requires more preliminary work (cf. [2, 4]) while the proof of the high relative accuracy seems like a more demanding task.

2 The Real and the Complex Falk-Langemeyer Method

Here we comparatively present the main formulas of the complex and real Falk-Langemeyer algorithms from [1, 3].

Let A and B be two n by n complex Hermitian (real symmetric) matrices. The complex (real) Falk–Langemeyer method solves the generalized eigenvalue problem $Ax = \lambda Bx$ by generating a sequence of "congruent" matrix pairs $(A^{(1)}, B^{(1)})$, $(A^{(2)}, B^{(2)}), \dots$ where $A^{(1)} = A$, $B^{(1)} = B$ and

$$A^{(k+1)} = F_k^* A^{(k)} F_k, \quad B^{(k+1)} = F_k^* B^{(k)} F_k, \quad \Big(A^{(k+1)} = F_k^T A^{(k)} F_k, \quad B^{(k+1)} = F_k^T B^{(k)} F_k, \Big) \quad k \geq 1.$$

Here F_k^* (F_k^T) denotes the Hermitian transpose (transpose) of F_k. The transformation matrices F_k are nonsingular *elementary plane matrices* with unit diagonal. Each F_k differs from the identity in only two elements at positions $(i(k), j(k))$ and $(j(k), i(k))$, where $1 \leq i(k) < j(k) \leq n$. The pair $(i(k), j(k))$ is called *pivot pair* and the 2×2 matrix $\hat{F}_k = [e_{i(k)}, e_{j(k)}]^* F_k [e_{i(k)}, e_{j(k)}]$ is called *pivot submatrix* of F_k. Here $e_1, \dots, e_n$ are the columns of the identity matrix I_n. For the CFL method we assume

$$\hat{F}_k = \begin{bmatrix} 1 & \alpha_k \\ \beta_k & 1 \end{bmatrix}, \quad \text{where } \alpha_k \text{ and } \beta_k \text{ are chosen to imply the condition } a_{i(k)j(k)}^{(k+1)} = 0\,, \quad b_{i(k)j(k)}^{(k+1)} = 0, \quad k \geq 1.$$

Here, $A^{(k)} = (a_{rt}^{(k)})$, $B^{(k)} = (b_{rt}^{(k)})$, $k \geq 1$. If the eigenvectors are wanted, the sequence of matrices $F^{(1)}$, $F^{(2)}$, ... has to be computed, where $F^{(1)} = I\,, \quad F^{(k+1)} = F^{(k)} F_k\,, \quad k \geq 1$. A way how the pivot pairs are selected is called *pivot*

* Corresponding author: e-mail hari@math.hr,

Copyright line will be provided by the publisher

strategy. The most common are the cyclic pivot strategies and among them the serial ones which annihilate the pivot elements $a^{(k)}_{i(k)j(k)}$ and $b^{(k)}_{i(k)j(k)}$ in the row- or column-wise fashion. The method is *globally convergent* if the scaled iterated matrices $A^{(k)}_S = [D^{(k)}]^{-1/2}A^{(k)}[D^{(k)}]^{-1/2}$ and $[D^{(k)}]^{-1/2}B^{(k)}[D^{(k)}]^{-1/2}$ tend to diagonal matrices. Here $D^{(k)}$ is the diagonal matrix whose diagonal elements are $\sqrt{(a^{(k)}_{rr})^2 + (b^{(k)}_{rr})^2}$, $1 \leq r \leq n$.

To describe the complex (real) Falk-Langemeyer algorithm we consider one step of the method. We drop the superscript k and denote the pivot pair by (i, j). Thus, $A = (a_{rt})$ and $B = (b_{rt})$ are the current Hermitian (symmetric) iteration matrices. Below are given the formulas for computing the transformation parameters α and β of the transformation matrix F (see [1, 3, 8]). They should be read in the order as they are written, from top to bottom. The formulas in parentheses refer to the real Falk-Langemeyer algorithm.

$$
\begin{aligned}
\Im_{ii} &= a_{ii}b_{ij} - a_{ij}b_{ii} = \begin{vmatrix} a_{ii} & b_{ii} \\ a_{ij} & b_{ij} \end{vmatrix} && \left(\Im_{ii} = a_{ii}b_{ij} - a_{ij}b_{ii} = \begin{vmatrix} a_{ii} & b_{ii} \\ a_{ij} & b_{ij} \end{vmatrix} \right) \\
\Im_{jj} &= a_{jj}b_{ij} - a_{ij}b_{jj} = \begin{vmatrix} a_{jj} & b_{jj} \\ a_{ij} & b_{ij} \end{vmatrix} && \left(\Im_{jj} = a_{jj}b_{ij} - a_{ij}b_{jj} = \begin{vmatrix} a_{jj} & b_{jj} \\ a_{ij} & b_{ij} \end{vmatrix} \right) \\
\Im_{ij} &= \Im'_{ij} + \imath\Im''_{ij},\ \Im'_{ij} = a_{ii}b_{jj} - a_{jj}b_{ii} = \begin{vmatrix} a_{ii} & b_{ii} \\ a_{jj} & b_{jj} \end{vmatrix} && \left(\Im_{ij} = a_{ii}b_{jj} - a_{jj}b_{ii} = \begin{vmatrix} a_{ii} & b_{ii} \\ a_{jj} & b_{jj} \end{vmatrix} \right) \\
\imath\Im''_{ij} &= a_{ij}\bar{b}_{ij} - \bar{a}_{ij}b_{ij} = \begin{vmatrix} a_{ij} & b_{ij} \\ \bar{a}_{ij} & \bar{b}_{ij} \end{vmatrix} = -2\imath \begin{vmatrix} \mathrm{Re}(a_{ij}) & \mathrm{Re}(b_{ij}) \\ \mathrm{Im}(a_{ij}) & \mathrm{Im}(b_{ij}) \end{vmatrix} \\
\Im &= \Im^2_{ij} + 4\bar{\Im}_{ii}\Im_{jj} = (\Im'_{ij})^2 - (\Im''_{ij})^2 + 2\imath\Im'_{ij}\Im''_{ij} + 4\bar{\Im}_{ii}\Im_{jj} && \left(\Im = \Im^2_{ij} + 4\Im_{ii}\Im_{jj} \right) \\
\nu &= (\Im_{ij} + \mathrm{sgn}(\Im'_{ij})\sqrt{\Im})/2, && \left(\nu = (\Im_{ij} + \mathrm{sgn}(\Im_{ij})\sqrt{\Im})/2 \right) \\
\alpha &= \Im_j/\nu, \quad \beta = -\bar{\Im}_i/\nu && \left(\alpha = \Im_j/\nu, \quad \beta = -\Im_i/\nu \right)
\end{aligned}
$$

If the pivot submatrices $\begin{bmatrix} a_{ii} & a_{ij} \\ \bar{a}_{ij} & a_{jj} \end{bmatrix}$ and $\begin{bmatrix} b_{ii} & b_{ij} \\ \bar{b}_{ij} & b_{jj} \end{bmatrix}$ are proportional, all the quantities $\Im_{ii}$, $\Im_{jj}$, $\Im_{ij}$, $\Im$ and ν are zero and a special algorithm is required. This is especially important when the computation is performed in finite arithmetic. Then all those quantities can be very tiny. Such a procedure is described in [3].

The remaining transformations that are part of one step of the method, which include the formulas how the diagonal and off-diagonal elements of A and B are updated, or how the columns of the matrix of accumulated transformations are updated, are quite easy to derive and are left to the reader.

We see that the new formulas for the complex Falk-Langemeyer algorithm are extensions of the existing formulas for the real algorithm. Hence we are confident that the proof of the asymptotic convergence from [8] and the accuracy estimates from [6] that are derived for the real method, can be also derived for the complex method. However, the most interesting and challenging research goal, which revealed itself during the numerical tests, is to prove the high relative accuracy of the real and complex method when the initial A and B are well-behaved positive definite matrices.

Acknowledgements This work has been fully supported by Croatian Science Foundation under the project IP-2014-09-3670.

References

[1] S. Falk, P. Langemeyer, Das Jacobische Rotations-Verfahren für reel symmetrische Matrizen-Paare I, II, Elektronische Datenverarbeitung, 30-43 (1960).

[2] V. Hari, E. Begović Kovač, *Convergence of the Cyclic and Quasi-cyclic Block Jacobi Methods*, Electronic transactions on numerical analysis **46**, 107-147 (2017).

[3] V. Hari, On Complex Falk-Langemeyer Method, proposed for publication in Numerical Algorithms.

[4] V. Hari, Globally convergent Jacobi methods for positive definite matrix pairs, Numerical Algorithms (to appear) https://doi.org/10.1007/s11075-017-0435-5.

[5] J. Matejaš, Accuracy of the Jacobi Method on Scaled Diagonally Dominant Symmetric Matrices. SIAM. J. Matrix Anal. Appl. **31(1)**, 133-153 (2009).

[6] J. Matejaš, Accuracy of one step of the Falk-Langemeyer method. Numerical Algorithms **68(4)**, 645-670 (2015).

[7] V. Novaković, S. Singer, S Singer, *Blocking and Parallelization of the Hari – Zimmermann Variant of the Falk – Langemeyer Algorithm for the Generalized SVD*, Parallel Comput. **49**, 136-152 (2015).

[8] I. Slapničar, V. Hari, On the Quadratic Convergence of the Falk-Langemeyer Method for Definite Matrix Pairs. SIAM J. Matrix Anal. Appl. **12(1)**, 84-114 (1991).

[9] K. Zimmermann, On the Convergence of the Jacobi Process for Ordinary and Generalized Eigenvalue Problems. Ph.D. Thesis, Dissertation No. 4305 ETH, Zũrich (1965).

Copyright line will be provided by the publisher